SOME EFFECTS OF CONTINUOUS TEMPO AND PITCH TRANSFORMATIONS IN PERCEIVED PLEASANTNESS OF LISTENING TO A MUSICAL SOUND FILE

Matti Gröhn, Lauri Ahonen, Minna Huotilainen

Finnish Institute of Occupational Health
Brain and Technology team
Topeliuksenkatu 41 a A, FIN-00250 Helsinki, Finland
`firstname.lastname@ttl.fi`

ABSTRACT

Physiological changes, i.e., changes in the heart rate and its variability, respiratory patterns etc., are induced by listening to music. Both the low-level acoustic features of the sound and the perceived pleasantness of the music contribute to the type and strength of the physiological changes. Thus, for studying the effects of the pleasantness, it would be important to keep the low-level acoustic features as similar as possible in the sound samples.

In this project we explored the effects of continuous tempo and pitch transformations in perceived pleasantness of listening to a musical sound file. Subjects evaluated pleasantness according to seven step Likert scale in which one is the most unpleasant and seven is the most pleasant. According to subjects' judgment the changes in tempo affects less to perceived pleasantness than changes in pitch.

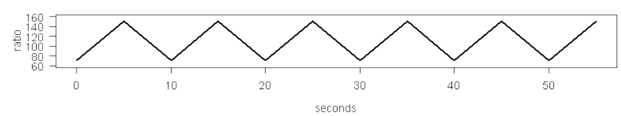
1. INTRODUCTION

Listening to music gives rise to strong reactions in the human brain and body, which are also seen as changes in different physiological parameters like heart rate and its variability, respiratory patterns, brain activity, skin conductance, muscular tension, etc. Positive mood can be induced by listening to the favourite music. In addition, also the low-level acoustic features of sounds have physiological effects. These include, e.g., rising of heart rate and blood pressure in the context of very loud sounds. In order to disentangle the effects of the low-level acoustic features and the subjective pleasantness, we need sound samples that share all relevant low-level features but differ in their pleasantness.

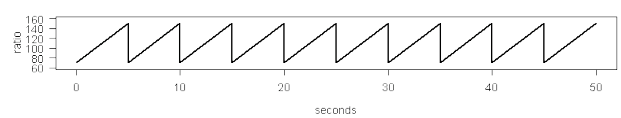
The listening experience is subjective and depends on the listener's musical taste, listening history, musical training, life events, the listening situation, and several other factors. Yet, there are several typical features that are found to be common in music that the listeners choose for a specific purpose, for example, for relaxation, even though their musical tastes vary. The musical features related to, for example, happy music include faster tempi, major keys, and high-pitched and ascending melodic lines, whereas sad music is associated with slow tempi, minor keys, and low-pitched and descending melodic lines.

When studying the physiological effects of music, it is important that the musical pieces used in the studies would have as similar sound properties as possible, in terms of mean tempi, frequency spectrum, mean loudness, etc. [1]. Simultaneously, when physiological effects related to pleasantness of music are studied, the actual pleasantness of the chosen musical excerpts should vary. For this reason, musical material evoking pleasant and unpleasant listening experiences but being similar in general sound properties is needed in order to disentangle the low-level physiological and pleasantness-related factors.

Our goal was to create modifications to music so that the physiologically relevant musical features would remain relatively constant while the features relevant to the listening experience would be modified. Our goal was that these modifications would significantly decrease the pleasantness of the chosen musical piece. As modifications, we decided to use local transformations of pitch and tempo (see section 3. Method). Our hypothesis was that the original musical piece without any transformations would be perceived as most pleasant by the listeners. Our second hypothesis was that, compared to the transformations of pitch or tempo alone, transformations in both dimensions simultaneously would produce the least pleasant listening experience.



Condition 1



Condition 2

Figure 1: Two different conditions to change tempo and pitch ratio.

2. RELATED RESEARCH

There are a number of experimental techniques that have been developed to induce positive or negative mood in the participants. The effectiveness of these Mood induction procedures (MIP) has been investigated [2] and the MIPs using music (with or without instruction, with or without visual material) have been shown to be effective. For example, the cardiovascular and respiratory patterns are changed according to the mood induced by music [3]. When studying the physiological features related to changes of mood, e.g., measures of heart rate or heart rate variability, blood pressure, etc., the acoustic and especially rhythmic content of the musical material plays a key role. One may expect effect related to temporal synchronization of the physiological functions and the musical material. Such synchronization, or entrainment of especially respiration to the temporal characteristics of the music used in the mood induction may completely override all physiological effects in the study [4]. In order to avoid this, we chose to use the same musical material in order to induce both positive and negative moods. To induce negative mood, the music needs to be modified so that it maintains its qualities with respect to synchronization (i.e., average tempo, average spectral content), but that it loses its positive qualities as a mood inducer.

A practical solution to compare sound files is the MIRtoolbox [5],[6],[7], developed at the University of Jyväskylä, which offers an integrated set of functions written in Matlab, dedicated to the extraction from audio files of musical features such as tonality, rhythm, structures, etc. Additionally to the basic computational processes, the toolbox also includes higher-level musical feature extraction tools. We used this tool in stimuli analysis as described in section 3. Method.

Fritz et al [8] have explored how a spectral manipulation of original, naturalistic music affects the perceived pleasantness of music. The spectral manipulation modified, among other factors, the sensory dissonance of music. Their results suggest that consonance and permanent sensory dissonance universally influence the perceived pleasantness of music.

3. METHOD

The task of the subjects were to scale perceived pleasantness according to seven step Likert scale, in which one was the most unpleasant and seven the most pleasant. They listen samples in quiet office room using Sennheiser HD 650 headphones.

3.1. Stimuli

We use six different stimuli in our experiment. One of them was the original 50 seconds musical excerpt (Pink Floyd, The Wall, Another Brick in the Wall Part 1). For the modified signals we used Adobe Audition 3.0 software for editing. For the tempo and pitch transformation we used two different conditions (see Fig. 1). In Condition 1, the ratio increase from 70 to 150 in five seconds, during the next five seconds the ratio decrease from 150 back to 70. In Condition 2, the ratio increase from 70 to 150 in five seconds and it repeated this pattern. For the tempo transformation we used Time Stretch effect, which preserves pitch. For the pitch transformation we used Pitch Shift effect, which preserves tempo. Finally, we used Pitch Bender effect to create signal in which both pitch and tempo were modified. After the transformation we applied 10 seconds fade out to each signal (including Original). In Table 1 are all the stimuli and their modification method. As seen in Fig. 2 waveforms were quite similar in modified signals than in original.

We used MIRtoolbox to analyze our stimuli set. In analysis we applied following MIRtoolbox functions [9].

- *mirdist* evaluates the distance between audio files along a particular representation specified by the user, corresponding to audio or musical features computed using MIRtoolbox.
- *mirchromagram* computes the chromagram of the stimulus. The chromagram is a redistribution of the spectrum energy along the different pitches (i.e., chromas):

Stimulus	Transformation method
Original	none
Tempo1	Tempo transformation, condition 1
Tempo2	Tempo transformation, condition 2
Pitch 1	Pitch transformation, condition 1
Pitch 2	Pitch transformation, condition 2
Both	Tempo and pitch transformation, spline

Table 1: Stimuli transformation methods.

Stimulus	Chromagram	Pulse clarity
Original	0.0000	0.0
Tempo1	0.0002	0.47
Tempo2	0.0001	0.50
Pitch 1	0.0046	0.02
Pitch 2	0.0031	0.08
Both	0.0027	0.40

Table 2: Chromagram and Pulse clarity distances to original stimulus according to MIRtoolbox analyses.

- *mirpulseclarity* [10] estimates the rhythmic clarity, indicating the strength of the beats estimated by the *mirtempo* function.
- *mirkeystrength* computes the key strength, i.e., the probability associated with each possible key candidate, through a cross-correlation of the chromagram returned by *mirchromagram*, wrapped and normalized (using the Normal option), with similar profiles representing all the possible tonality candidates.

In analysis we verified, that transformations only affect on transformed feature and left other features unaffected. In our approach, redistribution of the spectrum energy is important and transformations should not affect on it. As seen in Table 2, modified stimuli are very close to the Original in Chromagram distance. In the same Table, it is easy to see, that Tempo1, Tempo2 and Both stimuli differ from Original in Pulse clarity distance, but Pitch1 and Pitch2 are close to the original.

We used *mirkeystrength* function to analyze pitch transformation effects. Key strength value 1 means that music excerpt is certainly in that key, and -1 that is certainly not in that key. These values are provide for each twelve tonal keys. In Table 3 are minimum and maximum key strength values for each stimulus. Original, Tempo1 and Tempo2 stimuli have a clear key candidates. Pitch transformation has affected as expected and Pitch1, Pitch2 and Both stimuli don't have clear key candidates. In addition, in Fig. 3 Original, Tempo1 and Tempo2 stimuli have almost identical key strength graphs. On the other hand, key strength graphs for Pitch1, Pitch2 and Both stimuli are clearly different.

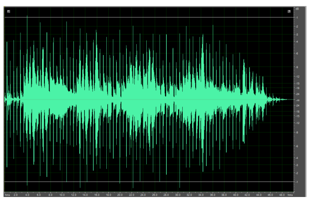
Our analysis verified, that transformations only affect on transformed feature and maintained other features almost unaffected.

3.2. Subjects

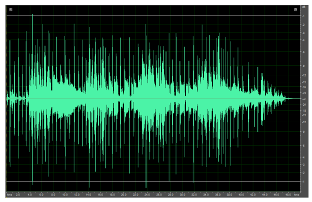
For this experiment we had 6 non-paid volunteers. Each of them reported to have normal hearing, although this was not verified with audiometric tests. We had 3 male subjects and 3 female subjects.

Stimulus	Key strength min	Key strength max
Original	-0.8	0.8
Tempo1	-0.8	0.8
Tempo2	-0.8	0.8
Pitch 1	-0.3	0.4
Pitch 2	-0.3	0.3
Both	-0.5	0.5

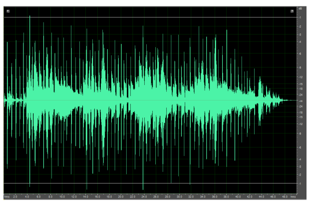
Table 3: Key strength minimum and maximum values according to MIRtoolbox analyses.



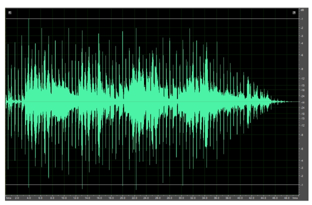
(a) Original



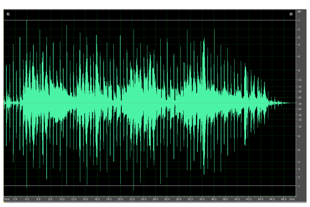
(b) Tempo1



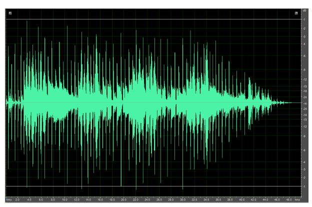
(c) Tempo2



(d) Pitch1



(e) Pitch2



(f) Both

Figure 2: Waveforms of the all the stimulus.

3.3. Procedure

Each subject first listened to the original signal. After that they listened to three sets of stimuli. Each set contained each stimulus once in a different randomized order and each subject listened to a different combination of these sets to avoid learning effects in results. Subjects evaluated the pleasantness of each stimulus immediately after listening to it.

4. RESULTS

In Table 4 are means for perceived pleasantness for each stimulus for all and for each subject. In Table 5 are medians for perceived pleasantness for each stimulus for all and for each subject. In general, the original stimulus was perceived as the most pleasant stimulus. The perceived pleasantness of Tempo1 and Tempo2 stimuli were better than other modified stimuli. The perceived pleasantness of Both stimulus was in the same level as Pitch1 and Pitch2 stimuli.

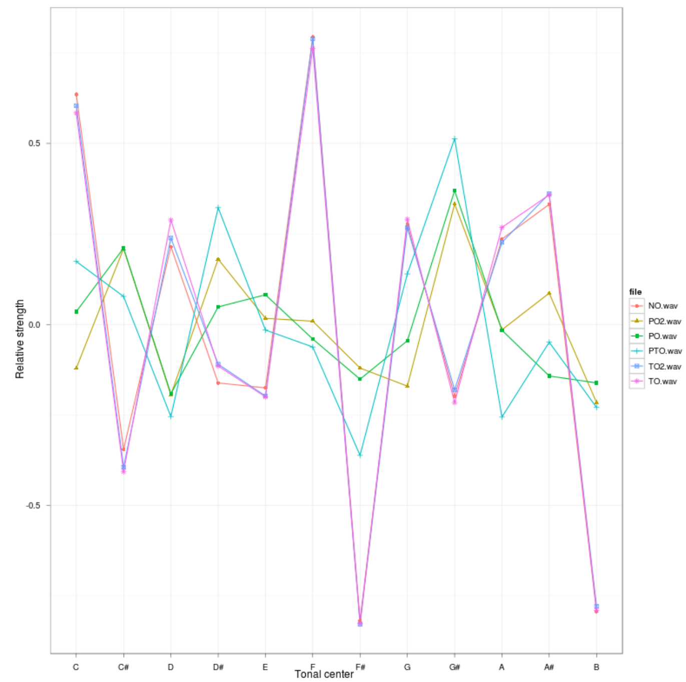


Figure 3: Key strength graphs of the all the stimulus. Original = NO, Tempo1 = TO, Tempo2 = TO2, Pitch1 = PO, Pitch2 = PO2 and Both = PTO

	Stimulus					
	Orig.	**Tempo1**	**Tempo2**	**Pitch1**	**Pitch2**	**Both**
s1	4.00	2.33	2.67	4.00	2.67	2.33
s2	5.67	3.67	3.33	2.33	1.67	2.00
s3	6.00	4.67	3.67	1.67	2.33	2.67
s4	7.00	5.00	4.67	2.00	3.00	2.33
s5	5.33	2.33	2.00	1.00	1.33	1.33
s6	6.00	4.00	4.00	3.33	2.00	4.00
s1-s6	5.67	3.67	3.39	2.39	2.17	2.44

Table 4: Mean ratings for all subjects.

In general, each subject evaluated stimuli-set in the same way. There were two exceptions. Subject 1 perceived Pitch1 stimulus as the most pleasant (see Table 5). The Subject 6 perceived Both stimulus as pleasant as Tempo1 and Tempo2 stimuli (see Table 5).

According to pairwise t-test (see Table 6), the perceived pleasantness of the original stimulus was statistically significantly better than modified stimuli. The difference between tempo modified (Tempo1 and Tempo2) and pitch modified (Pitch1 and Pitch2) stimuli was also statistically significant. There were no statistically significant difference between the conditions both in tempo modified and pitch modified stimuli. The perceived pleasantness of Both stimulus was statistically significantly different from Tempo1 and Tempo2, but it was not statistically different from Pitch1 and Pitch2.

According to our results the changes in tempo affects less to perceived pleasantness than changes in pitch as seen in Fig. 4.

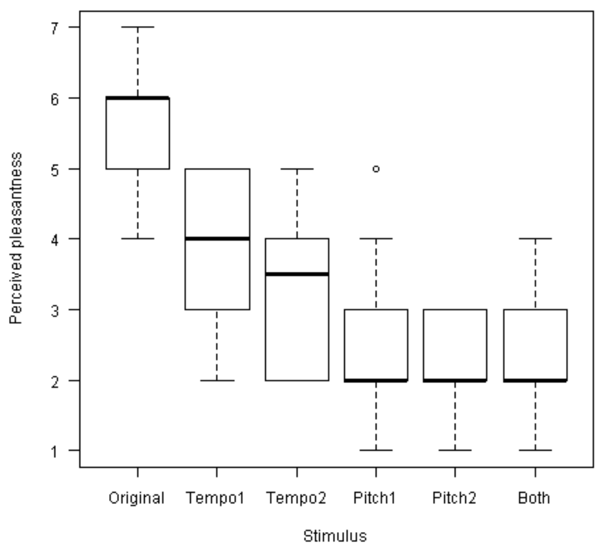


Figure 4: Boxplot distribution of stimulus ratings.

	Stimulus					
	Orig.	**Tempo1**	**Tempo2**	**Pitch1**	**Pitch2**	**Both**
s1	4.0	2.0	3.0	5.0	3.0	2.0
s2	6.0	4.0	3.0	2.0	2.0	2.0
s3	6.0	5.0	4.0	2.0	3.0	3.0
s4	7.0	5.0	5.0	2.0	3.0	2.0
s5	5.0	2.0	2.0	1.0	1.0	1.0
s6	6.0	4.0	4.0	3.0	2.0	4.0
s1-s6	6.0	4.0	3.5	2.0	2.0	2.0

Table 5: Median ratings for all subjects.

5. DISCUSSION

We found that it is possible to preserve the physiologically relevant musical features in music chosen by the listener while still reducing the pleasantness of the music. This is very relevant for future studies of physiological reactions to listening to music, since now it is possible to vary parametrically both the physiologically relevant sound features like intensity, tempo, etc, and separately pleasantness.

As expected in general subjects like the original signal most. This is the similar result as the Fritz et al [8] have achieved. Also in their experiment the original versions were preferred over modified versions. The Subject 1 informed us after the listening test, that she did not like Original stimulus much. That might explain, why she evaluated Pitch1 over the Original stimulus. Subject 6 told us after listening test, that she found the Both stimulus interesting. In addition, she evaluated it as pleasant as Tempo1 and Tempo2 stimuli. Although the general results are clear, the individual differences should be taken account, while further studies are planned and designed.

	Orig.	**Tempo1**	**Tempo2**	**Pitch1**	**Pitch2**
Tempo1	$p < 0.01$	-	-	-	-
Tempo2	$p < 0.01$	1.00	-	-	-
Pitch1	$p < 0.01$	$p < 0.01$	0.03	-	-
Pitch2	$p < 0.01$	$p < 0.01$	$p < 0.01$	1.00	-
Both	$p < 0.01$	$p < 0.01$	$p < 0.01$	1.00	1.00

Table 6: Pairwise t-test p-values for each stimulus pair.

We also found that transformations in pitch were found to be less pleasant than corresponding transformations in tempo. According to our results the effect of pitch transformations is so strong, that adding simultaneous tempo transformation does not decrease the perceived pleasantness. This may be related to the fact that in music, variations in tempo, albeit different from the ones used here, are used as an expressive element while variations in pitch produce continuous fluctuations of pitch leading to a perception of micro intervals, which is not typical for Western music. Thus, some of the tempo alterations could be interpreted as expressive (or over-expressive), while the pitch alterations can only be interpreted as a mistake of either the performer/performance or the equipment. These results are important for planning the modifications to be used in future studies of physiological reactions to music listening, since the effect of this modification is strong even though it still preserves all the important characteristics that mediate the physiological response to music. It is well known that listening to pleasant music has strong physiological effects mediated by the limbic system of the brain [11]. Also structural information of the musical piece is preserved in our modifications, which is important since structure of music has been shown to affect the emotional responses [12].

6. FUTURE RESEARCH

In future, it would be important to test the effects of these modifications on the pleasantness of different types of music. It is probable that some types of music are less or more vulnerable to these modifications in terms of pleasantness. Most probably, music that relies in its expression to simple, repetitive, continuous tempi and music where key clarity is strong is most strongly affected by these modifications.

In the future, these transformations can be used in mood induction experiments. In these studies, the pleasantness of the presented musical sounds can be varied by using the transformations presented here. Thus, the low-level acoustic features of the sounds are preserved, while the pleasantness varies. We plan to record changes in the physiology, brain activity, cognitive functions and mood of the participants prior to and after listening to musical soundfiles that are modified in the way presented in this paper.

7. ACKNOWLEDGMENT

We thank our subjects. In addition, we thank the reviewers for their valuable comments, which improved our paper.

8. REFERENCES

[1] D. Sammler, M. Grigutsch, T. Fritz, and S. Koelsch, "Music and emotion: Electrophysiological correlates of the pro-

cessing of pleasant and unpleasant music," *Psychophysiology*, vol. 44, pp. 293–604, 2007.

[2] R. Westermann, K. Spies, G. Stahl, and F. Hesse, "Relative effectiveness and validity of mood induction procedures: a meta-analysis," *European Journal of Social Psychology.*, vol. 34, pp. 557–580, 1992.

[3] J. Etzel, E. Johnsen, J. Dickerson, D. Tranel, and R. Adolphs, "Cardiovascular and respiratory responses during musical mood induction," *International Journal of Psychophysiology*, vol. 61, pp. 57–69, 2006.

[4] C. Wientjes, "Respiration in psychophysiology: methods and applications," *Biological Psychology.*, vol. 34, no. 2–3, pp. 179–203, 1992.

[5] O. Lartillot and P. Toiviainen, "A matlab toolbox for musical feature extraction from audio," in *Proc. of the 10th Int. Conference on Digital Audio Effects (DAFx-07)*, Bordeaux, France, 2007.

[6] https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox.

[7] O. Lartillot, P. Toiviainen, and T. Eerola, "A matlab toolbox for music information retrieval," in *Data Analysis, Machine Learning and Applications, Studies in Classification, Data Analysis, and Knowledge Organization*, C. Preisach, H. Burkhardt, L. Schmidt-Thieme, and R. Decker, Eds. Springer-Verlag, 2008.

[8] T. Fritz, S. Jentschke, N. Gosselin, D. Sammler, I. Peretz, R. Turner, A. Friederici, and S. Koelsch, "Universal recognition of three basic emotions in music," *Current Biology*, vol. 19, 2009.

[9] O. Lartillot, *MIRtoolbox 1.3.2, Users Manual*. Jyväskylä, Finland: Finnish Centre of Excellence in Interdisciplinary Music Research, University of Jyväskylä, 2011.

[10] O. Lartillot, T. Eerola, P. Toiviainen, and J. Fornari, "Multifeature modeling of pulse clarity: Design, validation, and optimization," in *Proc. of the International Conference on Music Information Retrieval*, Philadelphia, USA, 2008.

[11] A. Blood, R. Zatorre, P. Bermudez, and A. Evans, "Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions," *Nature Neuroscience*, vol. 2, pp. 382–387, 1999.

[12] J. Sloboda, "A. music structure and emotional response: some empirical findings," *Psychol.Music*, vol. 19, pp. 110–120, 1991.